

# One Forecast Engine

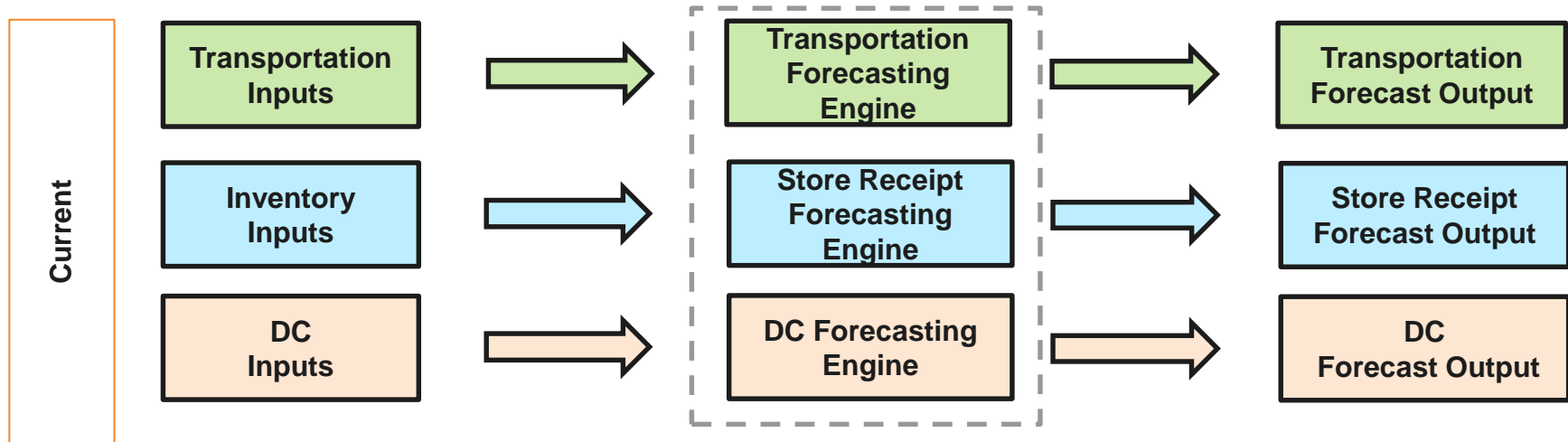
---

**Zach Cote, Data Scientist – SC Data Science**

**Kevin Kelleher, Data Scientist – SC Data Science**



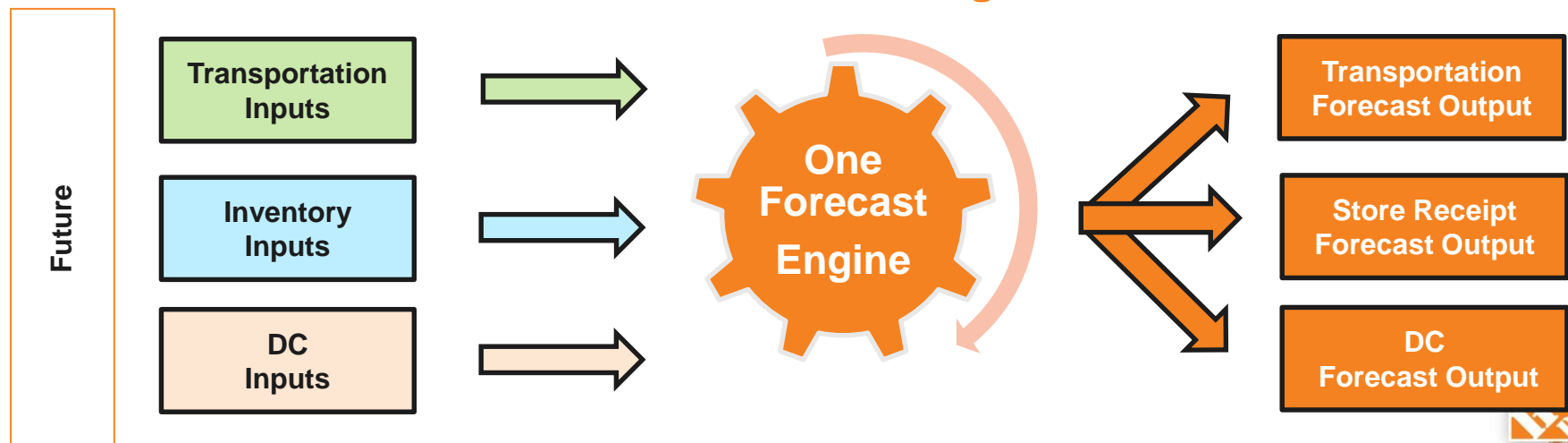
# One Forecast Engine – Vision



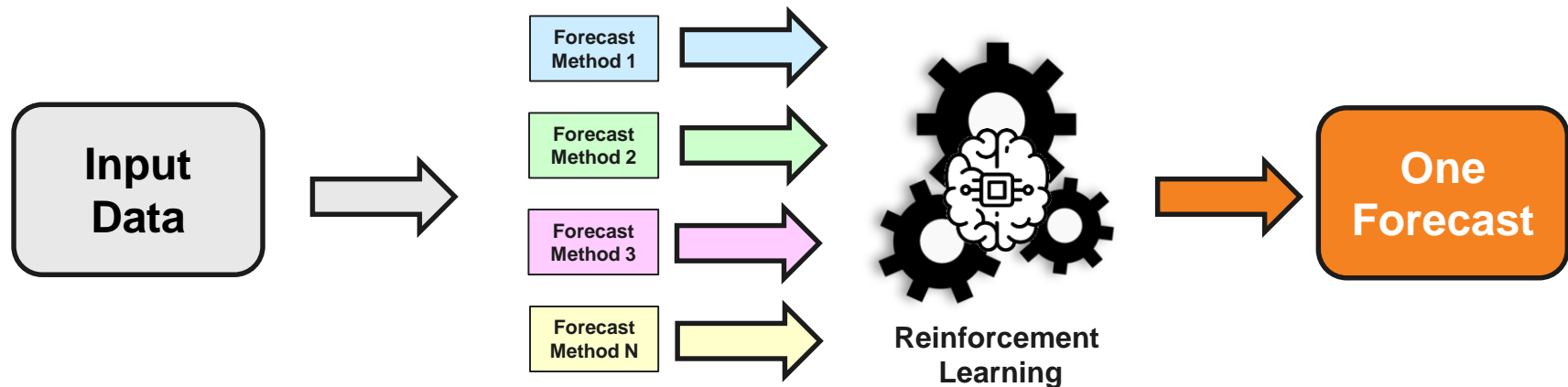
Unify separate forecasting engines into a universal engine

=

One Forecast Engine



# One Forecast Engine – Approach



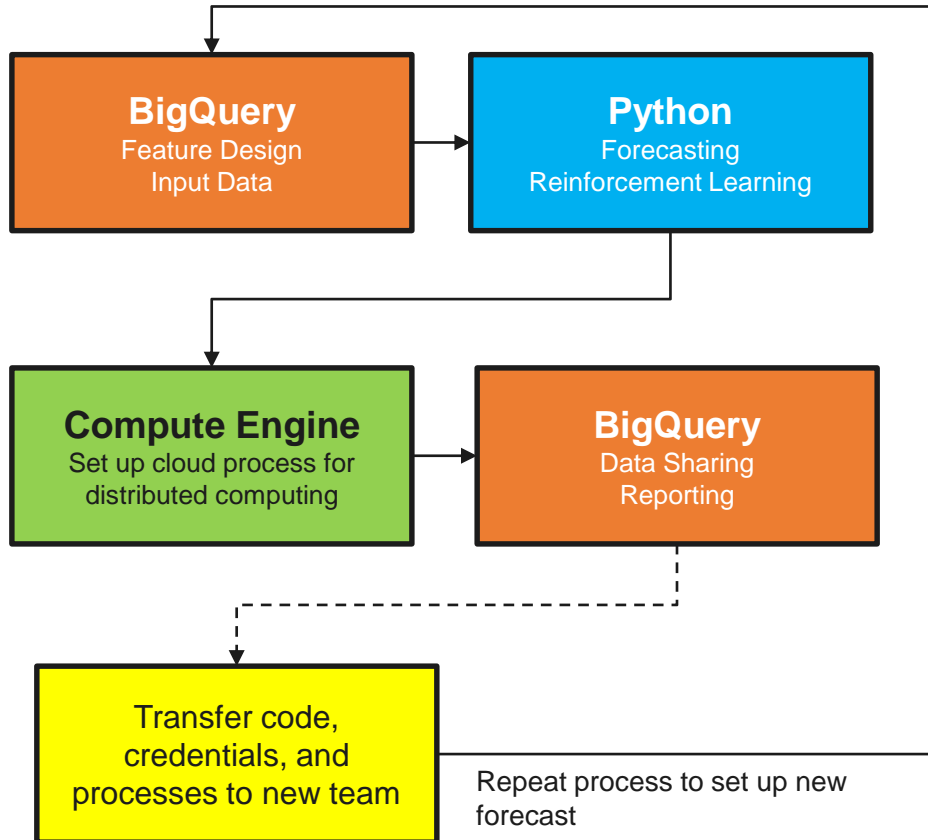
## ■ Ensemble Methodology:

- Leverage several distinct forecasting techniques to create multiple forecasts
- Continually evaluate which forecast technique produces the most accurate forecast
- Output a single forecast that outperforms each individual input forecast

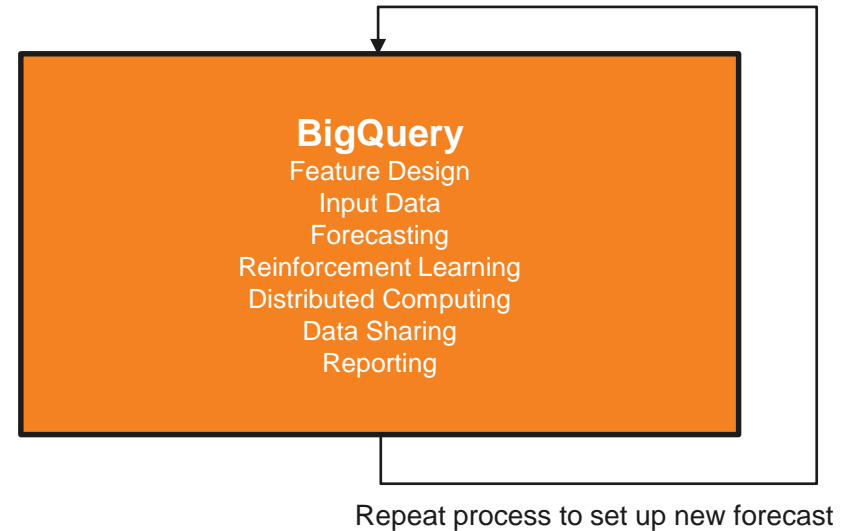


# Potential Architectures

## Traditional Architecture



## Our Architecture



# Benefits of Our Approach

---



Ease of use – generate multiple forecasts with a single button press



Runs quickly using already available cloud computing power



Complete integration with existing BigQuery data



Accessibility – empowers any analytics associate to create a forecast without previous forecasting experience



# Features - Overview

---

Forecast Creation

External Forecast Integration

Forecast Ensembling

Forecast Plotting

Forecast Adjustments

Code Exporting



# Features – Forecast Creation

## Forecast Creation Input:

Create up to three forecasts based on time series input

|                             |   |
|-----------------------------|---|
| Choose from preset options: | Weekly Lane Truckloads                          |
| Time Period:                | fscI_wk_end_dt                                  |
| Forecast Value Hierarchy:   | origin, destin                                  |
| Forecast Value:             | loads   |
| Forecast Periods:           | 13  |
| Dataset:                    | analytics-supplychain-thd.one_forecast          |
| Forecast Start:             | 02/04/2019                                      |
| Forecast End:               | 02/02/2020                                      |
| Training Flag:              | True  |
| Forecast Frequency:         | Week  |
| Input Data Table:           | `analytics-supplychain-thd.one_forecast.inputs` |

☐ Round Output

☐ Disallow Negative Outputs

Create Forecast

Clear Output

Cancel Run

Generate three distinct forecasts with one button click

Only requires one input table with time series data

## inputs

Schema Details [Preview](#)

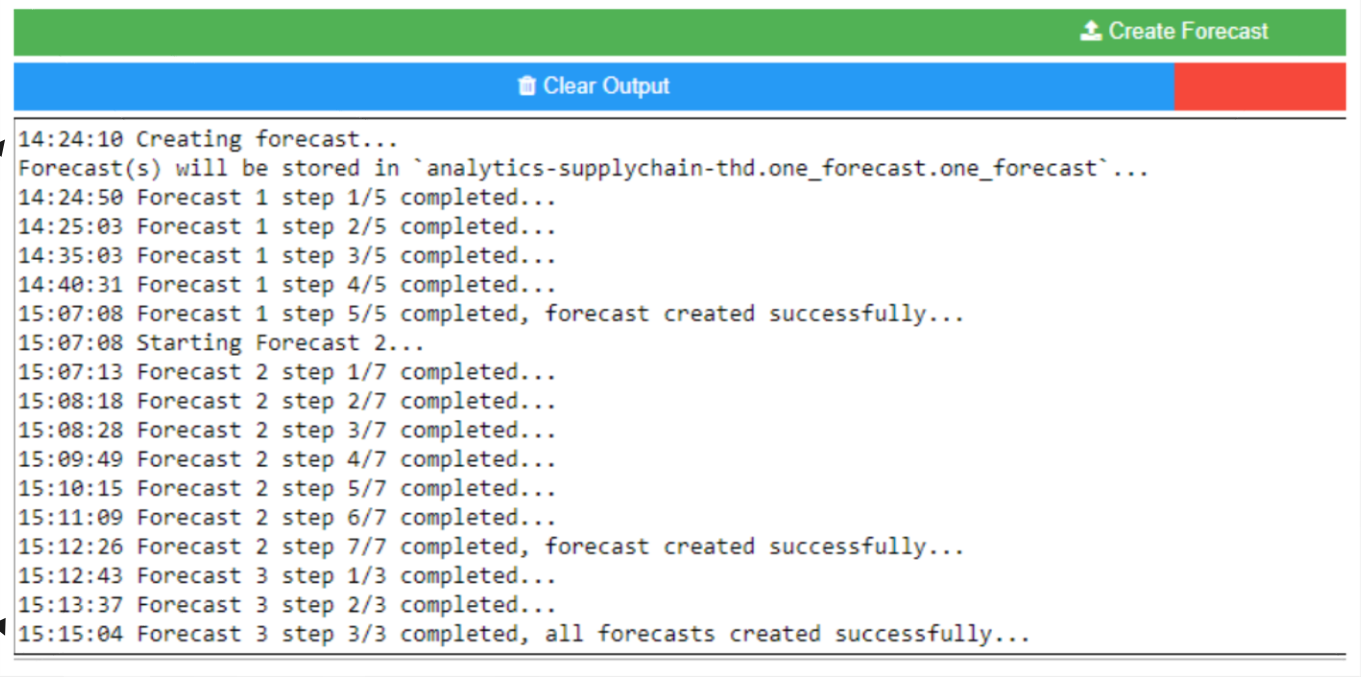
| Row | origin       | destin | fscI_wk_end_dt | loads |
|-----|--------------|--------|----------------|-------|
| 1   | 622101CA002  | 6006   | 2018-04-08     | 15.0  |
| 2   | 8746080H001  | 5781   | 2018-05-06     | 0.0   |
| 3   | 11165NJ001FB | 5982   | 2017-06-25     | 0.0   |
| 4   | 5621         | 5565   | 2014-08-17     | 29.0  |
| 5   | 5097         | 5920   | 2015-11-01     | 20.0  |
| 6   | 736100TX001  | 5852   | 2017-03-19     | 42.0  |



# Features – Forecast Creation

- Step by step output walking user through creation process
- Creates one year of 13-week snapshots for three forecasts in < 1 hour (based on 3.1M rows of input data)

Generates 3 forecasts  
in < 1 hour



```
14:24:10 Creating forecast...
Forecast(s) will be stored in `analytics-supplychain-thd.one_forecast.one_forecast`...
14:24:50 Forecast 1 step 1/5 completed...
14:25:03 Forecast 1 step 2/5 completed...
14:35:03 Forecast 1 step 3/5 completed...
14:40:31 Forecast 1 step 4/5 completed...
15:07:08 Forecast 1 step 5/5 completed, forecast created successfully...
15:07:08 Starting Forecast 2...
15:07:13 Forecast 2 step 1/7 completed...
15:08:18 Forecast 2 step 2/7 completed...
15:08:28 Forecast 2 step 3/7 completed...
15:09:49 Forecast 2 step 4/7 completed...
15:10:15 Forecast 2 step 5/7 completed...
15:11:09 Forecast 2 step 6/7 completed...
15:12:26 Forecast 2 step 7/7 completed, forecast created successfully...
15:12:43 Forecast 3 step 1/3 completed...
15:13:37 Forecast 3 step 2/3 completed...
15:15:04 Forecast 3 step 3/3 completed, all forecasts created successfully...
```





# Features – Integration with Existing Forecasts

- Upload an existing forecast to be compared to and/or ensembled with forecasts created by the tool

**External Forecast Input (Optional):**  
Add an external forecast to the one\_forecast table to be considered during the ensemble approach

|                               |   |
|-------------------------------|---|
| Forecast Value Hierarchy:     | List of column names to aggregate forecast at (ex: dept, class, subclass or origin, destination)                |
| Dataset:                      | Dataset to Output Forecast to (Must have a copy of previously generated one_forecast table)                     |
| Forecast Created Time Period: | Name of create date column in input query   |
| Forecast Time Period:         | Name of forecast date column in input query   |
| Forecast Value:               | Target value from forecast input  |
| Forecast Input Query:         | External Forecast Query (Must have columns representing Forecast Create Time, Forecast Time and Forecast Value) |

Insert Additional Forecast



# Features – Ensemble Method


- Combine previously created/uploaded forecasts into one forecast using AI that outperforms each individual input forecast
- Option to use rules-based selection instead

Uses JavaScript code stored in cloud storage bucket

**Ensemble:**  
Combine previously made forecasts into the best possible forecast

Forecast Selection Logic:

Cloud Storage Bucket Path:



```
09:33:49 Creating Selection Logic Input Tables...
09:33:49 Running chosen selection logic, results will be stored in `analytics-supplychain-thd.one_forecast.one_forecast`...
09:34:27 Selection logic input table 1/2 completed...
09:45:30 Selection logic input table 2/2 completed...
09:52:59 Reinforcement Learning selection output stored in `analytics-supplychain-thd.one_forecast.one_forecast`. Forecast process complete.
```



# Features – Built in Forecast Analysis

- In tool error analysis for all forecasts

## Forecast Analysis:

Plot newly created forecasts and error

Forecast Lag to Plot:

0

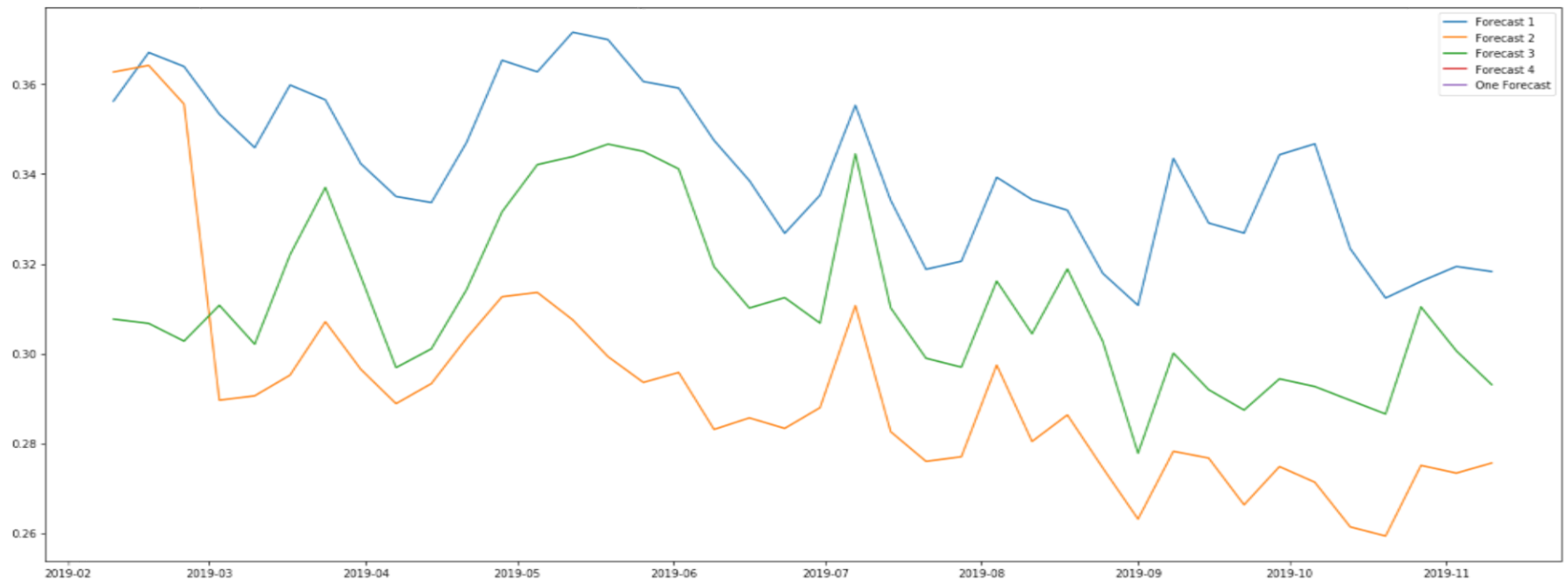
Plot Period Start:

12/10/2018

Plot Period End:

12/10/2019

Plot Forecast MAPE



# Features – Built in Forecast Analysis

- In tool forecast vs actual plotting



# Features – Adjust Forecast

- Options to adjust previously created forecasts with the following:
  - Corrections for consistent over/under forecasting
  - Overall forecast % lift/drop

## Forecast Adjustments:

Make adjustments to the recently created forecast

Forecast Adjustment Options

Forecast Increase - Across the board increase in forecast values by x%

Adjustment Periods

Number of periods to adjust off of (n)

Adjustment Percentage

3

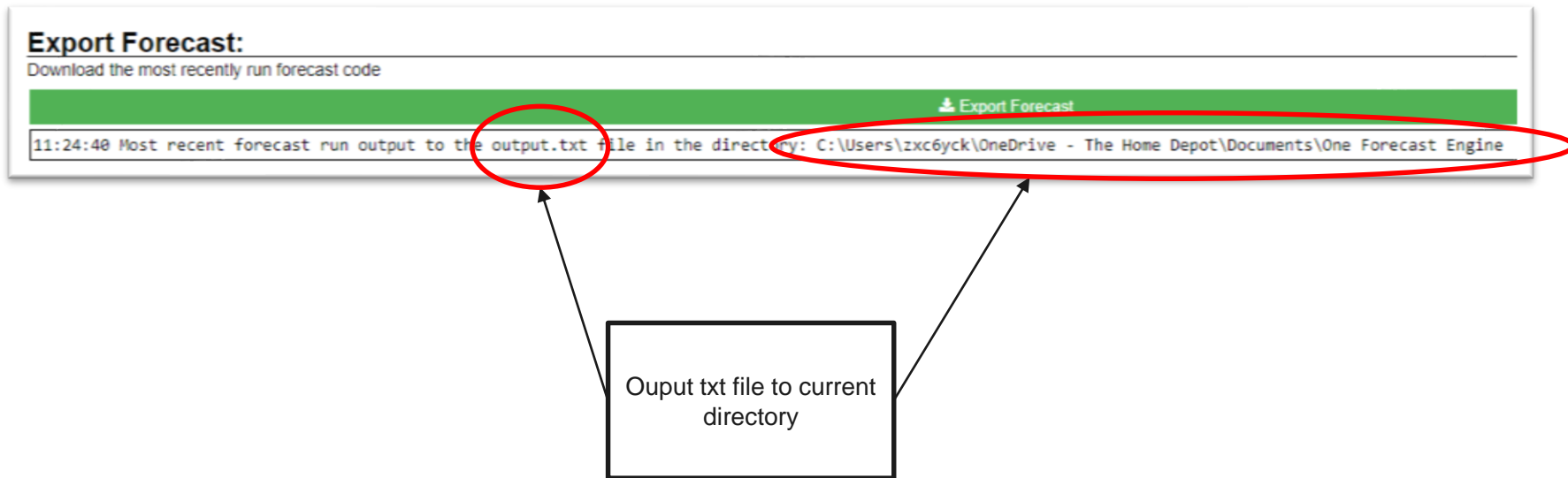
⚙ Adjust Forecast

16:38:52 Forecasts stored in `analytics-supplychain-thd.one\_forecast.one\_forecast` will be adjusted accordingly...  
16:39:32 Forecasts in `analytics-supplychain-thd.one\_forecast.one\_forecast` increased by 3%.

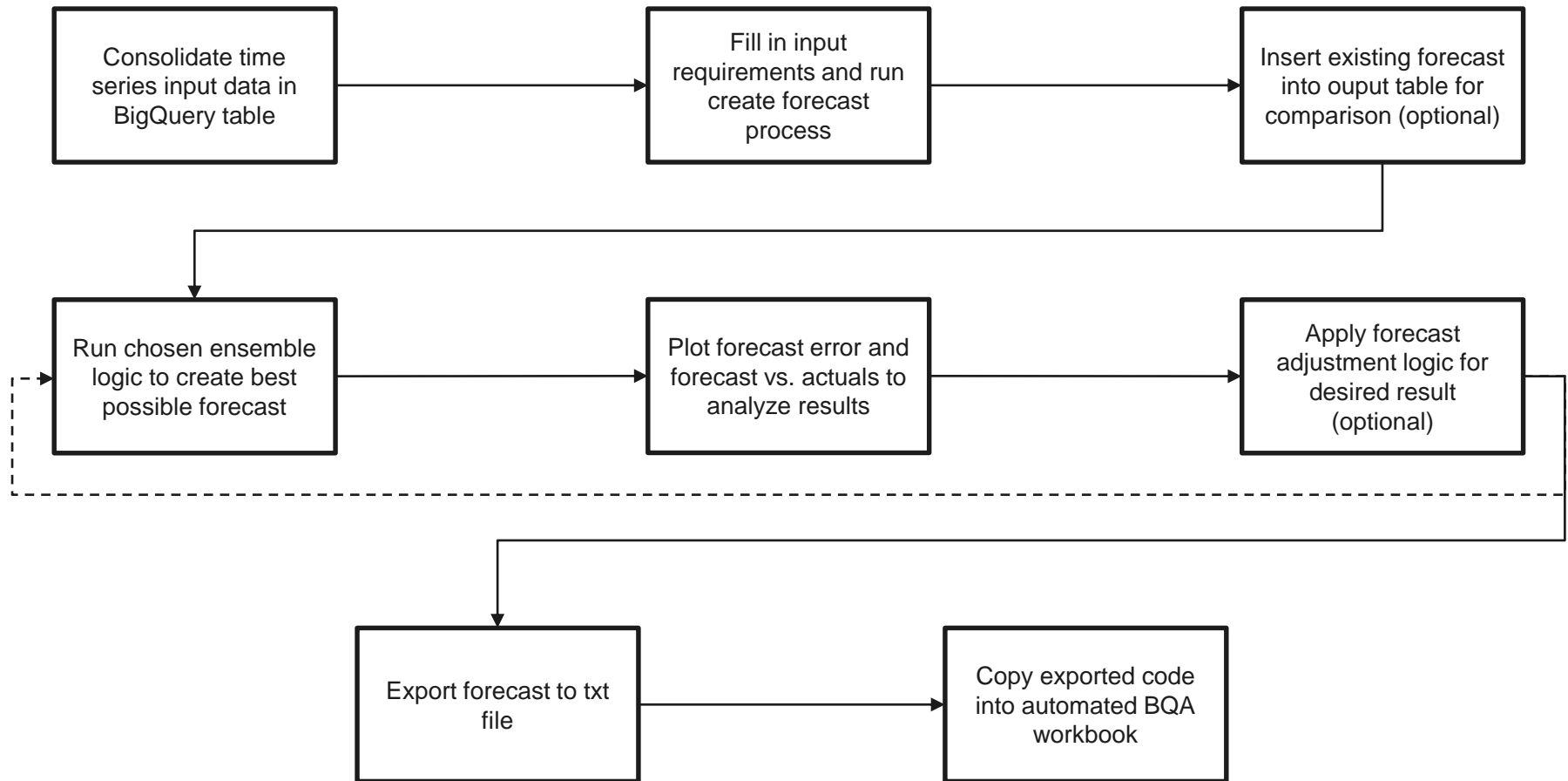


# Features – Export Forecast

- Output all code run in the current session to a txt file in the current directory
- Allows easy setup of an automated BQA workbook using our process



# Tool Workflow



# Installation

---





# Installation – Base Python Version

---

- Use this installation method to use the tool with your base python version
- The tool is run completely through a Jupyter Notebook which will require the following:
  - Anaconda for Python 3.7  
(<https://www.anaconda.com/distribution/#windows>)
  - Google-cloud-bigquery (pip install google-cloud-bigquery)
    - If there is an error with this package use pip install google-cloud-bigquery==1.21.0
  - Pandas-gbq (pip install pandas-gbq)
  - The Google Cloud SDK (outlined in the following slides)



# Installation – Custom Anaconda Environment

---

- Use this installation procedure if you want to set up a custom working environment for the tool
- Install Anaconda for Python 3.7  
((<https://www.anaconda.com/distribution/#windows>)
- Install the Google Cloud SDK (outlined in the following slides)
- Run the following code in the anaconda prompt
  - `conda create -n one_forecast python=3.7 google-cloud-bigquery=1.21.0 pandas=0.25.1 pandas-gbq=0.11.0 matplotlib=3.1.1 ipython=7.8.0 ipywidgets=7.5.1 -y --channel conda-forge`
  - `conda activate one_forecast`
  - `python -m ipykernel install --user --name=one_forecast`
- Then make the following selection in the Jupyter menu:
  - `kernel -> change kernel -> select the env named one_forecast`



# Download Cloud SDK

- This contains all the tools for the Cloud Platform, including BigQuery
- <https://cloud.google.com/sdk/docs/downloads-versioned-archives>
- Download the Windows 64-bit with Python Bundled zip file

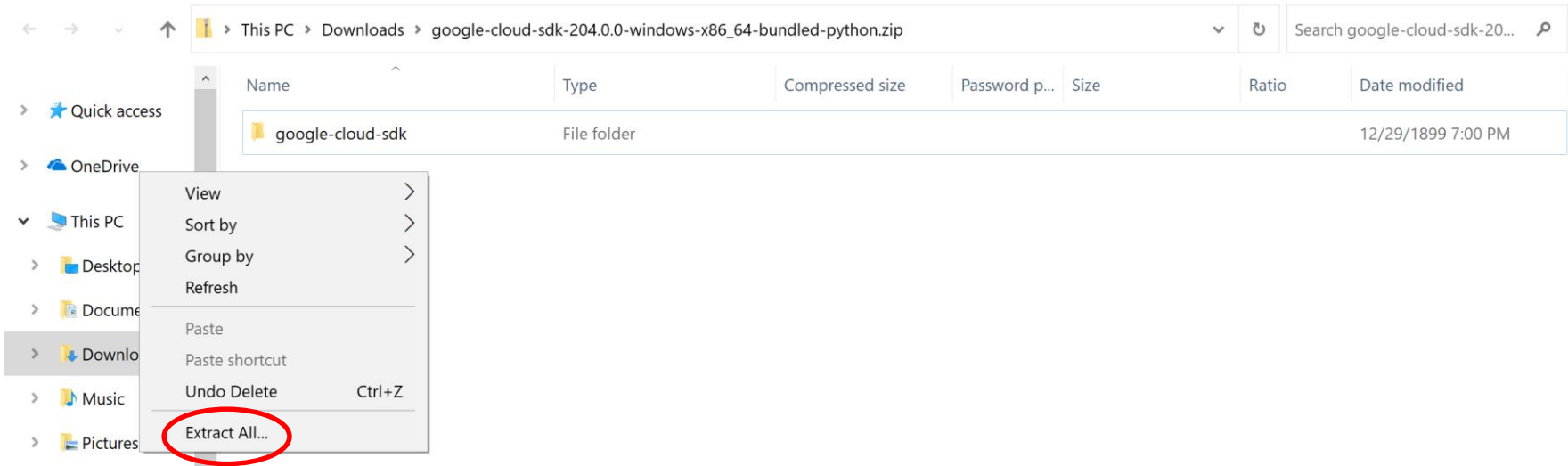


| Platform                                    | Package  | Size     | SHA256 Checksum  |
|---|--|----------|--|
| Linux 64-bit (x86_64)                       | <a href="#">google-cloud-sdk-204.0.0-linux-x86_64.tar.gz</a>               | 20.0 MB  | 276984a44a2a9dc1af5d3c859a1295897fd8cfc911738874daf007ab46143da5 |
| Linux 32-bit (x86)                          | <a href="#">google-cloud-sdk-204.0.0-linux-x86.tar.gz</a>                  | 19.6 MB  | 84737b674dee4e7ebb893be0ef2feaa869c98340069b6b8e9b3fa64834e0b339 |
| Mac OS X 64-bit (x86_64)                    | <a href="#">google-cloud-sdk-204.0.0-darwin-x86_64.tar.gz</a>              | 16.0 MB  | eaeaa9babf8e6c2a66bf6db3a2ecb34fc24fc4b2858ab3f7660386c7c79177cf |
| Mac OS X 32-bit (x86)                       | <a href="#">google-cloud-sdk-204.0.0-darwin-x86.tar.gz</a>                 | 16.0 MB  | 7bba1979ead36ef35a681024560c9edd822c7b53e923adee658af2cba8616dce |
| Windows 64-bit (x86_64)                     | <a href="#">google-cloud-sdk-204.0.0-windows-x86_64.zip</a>                | 103.4 MB | 5b9e38d37c983076cdad300e8df4b4a6c2dee3a898649d435e0cab92929656bc |
| Windows 64-bit (x86_64) with Python bundled | <a href="#">google-cloud-sdk-204.0.0-windows-x86_64-bundled-python.zip</a> | 141.8 MB | 86f4c33ef0ca26f88e9cd1658e513b799f7d7169e4e57e1e2cba57d1b074798a |
| Windows 32-bit (x86)                        | <a href="#">google-cloud-sdk-204.0.0-windows-x86.zip</a>                   | 103.6 MB | b110e40a49ae3ca6d3627aabb5e3fec34dbb0ac1f154cc854253a0f90b4a6da1 |
| Windows 32-bit (x86) with Python bundled    | <a href="#">google-cloud-sdk-204.0.0-windows-x86-bundled-python.zip</a>    | 137.9 MB | 230568ce0003c94a7ea39ce63bf890d323765f37d553d9e50dbacd907d41b331 |



# Install Cloud SDK

- Open the downloaded zip file:



- Right Click anywhere in this folder and **Extract All** to this path:
- C:\Users\<your LDAP here>\AppData\Local\Google\Cloud SDK

← Extract Compressed (Zipped) Folders

## Select a Destination and Extract Files

Files will be extracted to this folder:

C:\Users\<your LDAP here>\AppData\Local\Google\Cloud SDK

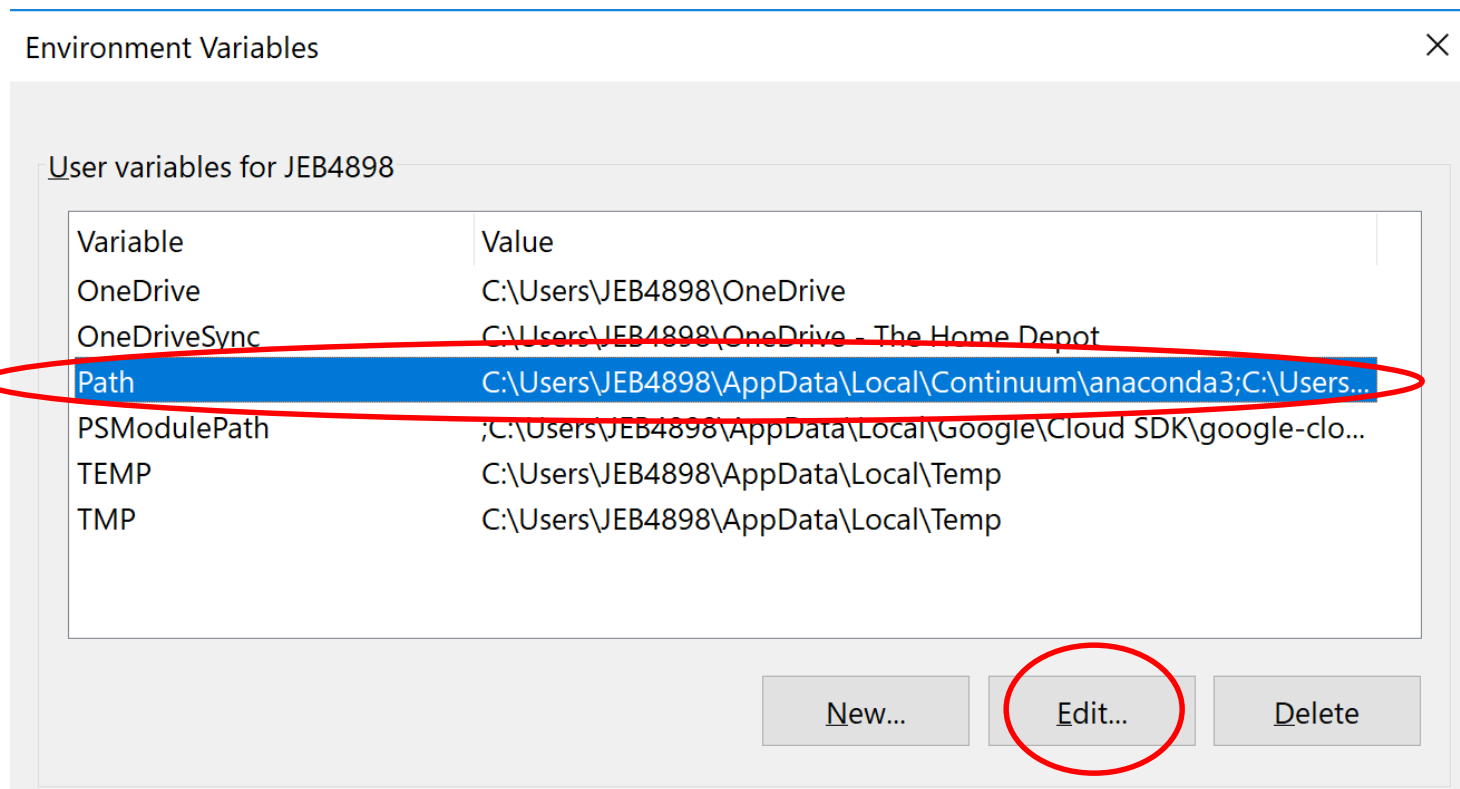
Browse...

☒ Show extracted files when complete



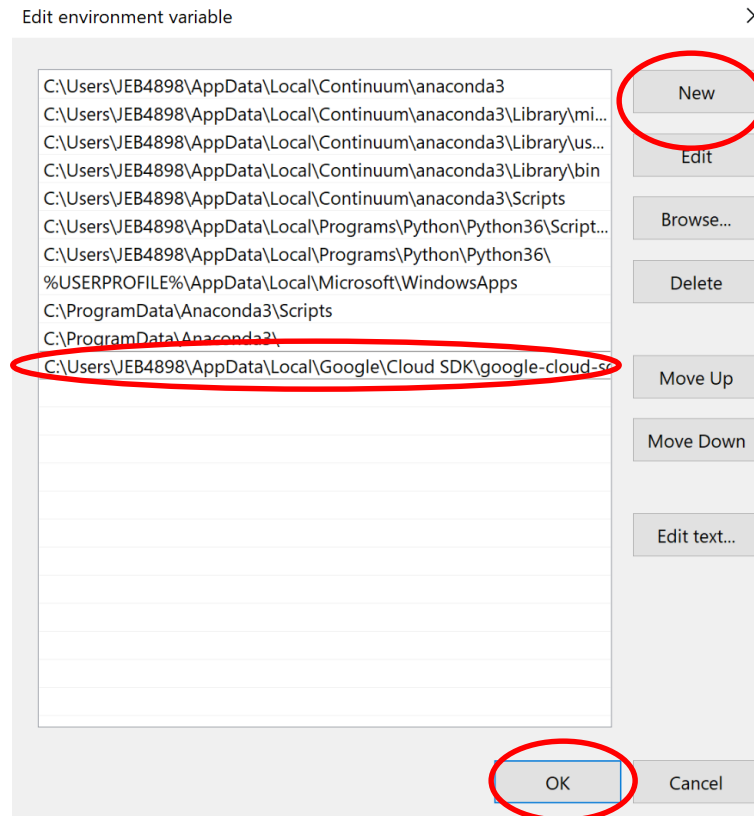
# Install Cloud SDK

- Go to your Environmental Variables, where we will edit the Path for gcloud



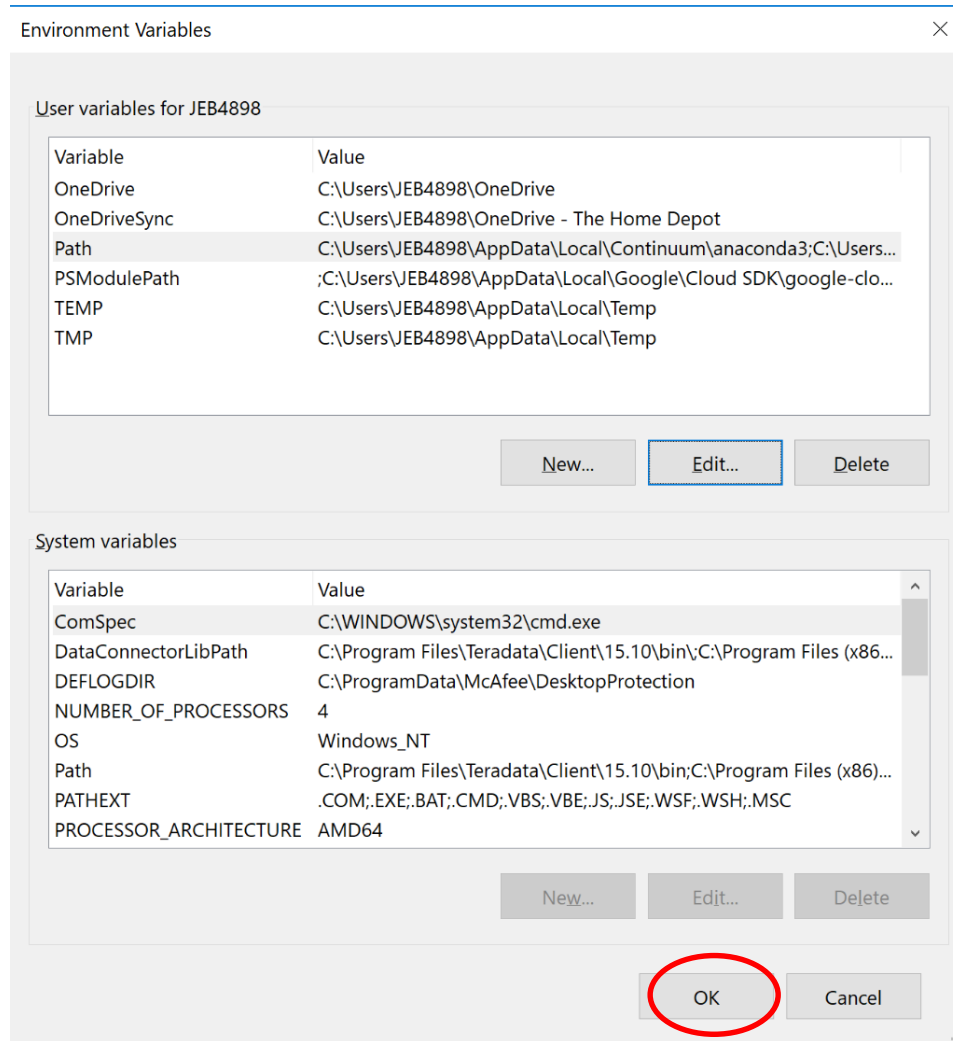
# Install Cloud SDK

- Click **New** to add a new variable with the same location you extracted your files to
- Type in the location:
  - C:\Users\<your LDAP here>\AppData\Local\Google\Cloud SDK\google-cloud-sdk\bin
- Click **OK**



# Install Cloud SDK

- Hit **OK** again to save the variable



# Install Cloud SDK

---

- In order to bypass the security issue that prevents you from finishing installation, open Command Prompt and run:

**\$ gcloud init --skip-diagnostics**

- You will receive this message: type in **Y** and hit enter

```
To continue, you must log in. Would you like to log in (Y/n)? Y
```

- When prompted, log into your account and allow Google Cloud SDK to have access
- Set your default project:
- Ours is **analytics-supplychain-thd**

```
Pick cloud project to use:
```

```
[1] [my-project-1]
```

```
[2] [my-project-2]
```

```
...
```

```
Please enter your numeric choice:
```





# Install Cloud SDK


- After initialization is complete:

**\$ gcloud auth application-default login**

```
C:\Users\JEB4898>gcloud auth application-default login
Your browser has been opened to visit:
```

```
https://accounts.google.com/o/oauth2/auth?redirect_uri=http%3A%2F%2Flocalhost%3A8085%2F&prompt=select_account&response_type=code&client_id=764086051850-6qr4p6gpi6hn506pt8ejuq83di341hur.apps.googleusercontent.com&scope=https%3A%2F%2Fwww.googleapis.com%2Fauth%2Fuserinfo.email+https%3A%2F%2Fwww.googleapis.com%2Fauth%2Fcloud-platform&access_type=offline
```

- Follow the prompts until you see:



Why Google Products Solutions Launcher Pricing Security Customers Documentation Support Partners

Cloud SDK

You are now authenticated with the Google Cloud SDK!

☆☆☆☆☆  
[SEND FEEDBACK](#)

The authentication flow has completed successfully. You may close this window, or check out the resources below.

Information about command-line tools and client libraries

To learn more about `gcloud` command-line commands, see the [gcloud Tool Guide](#).

For further information about the command-line tools for Google App Engine, Compute Engine, Cloud Storage, BigQuery, Cloud SQL and Cloud DNS (which are all bundled with Cloud SDK), see [Accessing Services with gcloud](#).

If you are a client application developer and want to find out more about accessing Google Cloud Platform services with a programming language or framework, see [Google APIs Client Libraries](#).

Tutorials

Here are some links to help you get started with Google Cloud Platform services.

